Kin-Cohort Designs for Gene Characterization

Mitchell H. Gail, David Pee, Raymond Carroll

Background: In the kin-cohort design, a volunteer with or without disease (the proband) agrees to be genotyped, and one obtains information on the history of a disease in first-degree relatives of the proband. From these data, one can estimate the penetrance of an autosomal dominant gene, and this technique has been used to estimate the probability that Ashkenazi Jewish women with specific mutations of BRCA1 or BRCA2 will develop breast cancer. *Methods:* We review the advantages and disadvantages of the kin-cohort design and focus on dichotomous outcomes, although a few results on time-to-disease onset are presented. We also examine the effects of violations of assumptions on estimates of penetrance. We consider selection bias from preferential sampling of probands with heavily affected families, misclassification of the disease status of relatives, violation of Hardy–Weinberg equilibrium, violation of the assumption that family members' phenotypes are conditionally independent given their genotypes, and samples that are too small to ensure validity of asymptotic methods. *Results and Conclusions:* The kin-cohort design has several practical advantages, including comparatively rapid execution, modest reductions in required sample sizes compared with cohort or case–control designs, and the ability to study the effects of an autosomal dominant mutation on several disease outcomes. The design is, however, subject to several biases, including the following: selection bias that arises if a proband's tendency to participate depends on the disease status of relatives, information bias from inability of the proband to recall the disease histories of relatives accurately, and biases that arise in the analysis if the conditional independence assumption is invalid or if samples are too small to justify standard asymptotic approaches. [Monogr Natl Cancer Inst 1999;26: 55–60]

Wacholder et al. *(1)* used the term "kin-cohort" design to describe a study in which volunteers (probands) are genotyped and one also determines the phenotypes of probands' first-degree relatives. This design was employed by Struewing et al. *(2)* to estimate the cumulative probability of developing breast cancer, as a function of age, for carriers of mutations of BRCA1 or BRCA2 in Ashkenazi Jews from the region surrounding Washington, DC. That study indicated that the cumulative risk of breast cancer to age 70 years was 0.56, a number that was substantially lower than estimates from highly affected pedigrees *(3),* but not inconsistent with the estimate of 0.69 obtained by Whittemore et al. *(4)* from segregation analyses (without any genotyping) of U.S. population-based case subjects and control subjects and their first-degree relatives. Although the term "kin-cohort design" might be used for any study of a family that is selected on the basis of a proband, we shall reserve its use to those studies in which the probands are genotyped.

For population-based inference, Gail et al. *(5)* stressed the importance of representative sampling of probands, conditional on their phenotypes. Thus, for dichotomous phenotypes with Y_0 = 1 denoting a case proband and Y_0 = 0 denoting a control proband, Gail et al. assumed that the case probands were selected at random from among all case subjects in the population and the control probands were selected at random from among all control subjects in the population. One can enrich the frequency of mutation carriers in a kin-cohort study by increasing the proportion of case probands.

Wacholder et al. *(1)* estimated the cumulative probability of developing breast cancer to a given age by recognizing that the cumulative incidence distribution for first-degree relatives of probands who carried the mutation was a mixture (nearly 50 : 50) of carrier and noncarrier distributions, whereas the cumulative incidence distribution for first-degree relatives of probands who did not carry the mutation was a different mixture (almost 0 : 100 of carrier and noncarrier distributions). It was thus possible to estimate the underlying carrier and noncarrier distributions from data on the cumulative incidence in relatives of carrier and non-carrier probands. Gail et al. *(5)* estimated penetrance using likelihood methods similar to those in segregation analysis but adapted to take into account the genotypes of the probands.

The kin-cohort design can be extended by genotyping some of the relatives of the proband *(5)* or by expanding the information on relatives to include the phenotypes of second-degree and more distant relatives. Gail et al. *(5)* used the term "genotyped-proband design" to stress that the proband is genotyped in the kin-cohort design. Studies of first-degree relatives of case and control probands in the absence of genotyping, such as those described by Whittemore et al. *(4)* and by Claus et al. *(6),* have been called "case–control family studies." These studies can be analyzed with the use of classical methods of segregation analysis. Thomas *(7)* used the term "family cohort study" whether or not the proband is genotyped.

There are several practical advantages of the kin-cohort design that invite its use for estimating penetrance *(1).* Often one can field and complete the study quickly. Because the disease status of the relatives can be determined at the date of recruitment of the proband, one does not need to wait to acquire such information as in a prospective cohort study. In some populations and for some diseases, one may obtain reliable information on the phenotypes of first-degree relatives simply by interviewing the proband. In such circumstances, time is saved because a single proband interview yields the phenotypes of several relatives. Moreover, one can inquire about more than one disease in the same proband interview. For example, Struewing et al. *(2)* demonstrated an increased risk for ovarian and prostate cancers among carriers of BRCA1 and BRCA2 mutations as well as an

Affiliations of authors: M. H. Gail, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD; D. Pee, Information Management Services, Rockville, MD; R. Carroll, Texas A&M University, College Station.

Correspondence to: Mitchell H. Gail, M.D., Ph.D., National Cancer Institute, EPS-8032, Bethesda, MD 20892 (e-mail: gailm@exchange.nih.gov).

See "Notes" following "References."

increased risk of breast cancer. One can also study the penetrances associated with additional mutations simply by genotyping the proband for the several mutations. Furthermore, as shown by Gail et al. *(5)*, the sample sizes required by the kin-cohort design to estimate penetrance of an autosomal dominant mutation with a desired degree of precision are comparable to or somewhat smaller than the sample sizes needed for a cohort study or for a population-based case–control design and can be substantially smaller if some relatives are also genotyped.

Estimates of penetrance from the kin-cohort design are, however, subject to certain biases that arise when assumptions underlying the analysis are not met. In this article, we review the assumptions underlying a likelihood-based analysis of the kin-cohort design (*see* "Methods" section) and give an example of sample sizes required by this design, compared with cohort or case–control designs (*see* "Results" section). Then we discuss the effects of violations of the assumptions on estimates of penetrance and allele frequency (*see* "Results" section). Finally, we consider some conditions favorable for the use of the kin-cohort design and some potential applications apart from estimating penetrance (*see* "Discussion" section).

METHODS

Several assumptions underlie a standard likelihood analysis of kin-cohort data, as described by Gail et al. *(5)*. For this article, we usually assume that phenotype is dichotomous, although methods for time-to-disease onset data have also been presented by Wacholder et al. *(1)* and Gail et al. *(5)*. A few simulations and analyses are based on parametric methods for survival data given by Gail et al. *(5)*, and we comment further on survival methods in the "Results" section. The key assumptions used in this article, some of which can be relaxed, are as follows: A1) Risk follows an autosomal dominant pattern, in which carriers of the mutant allele have a chance of disease (penetrance) ϕ_1 and noncarriers have penetrance ϕ_0; A2) the mutant allele "A" and wild-type "a" are in Hardy–Weinberg equilibrium (HWE); A3) given a family member's genotype, his or her phenotype is conditionally independent of the phenotypes of other family members; A4) probands with phenotype Y_0 are selected at random from members of the population with phenotype Y_0; A5) disease status is determined without error; and A6) sample sizes are large enough to justify standard asymptotic methods.

Under these assumptions, the likelihood for a typical family selected on the basis of a proband with phenotype Y_0 is

$$P(Y_1|g_0)P(g_0)|Y_0), \qquad [1]$$

where $g_0 = 1$ or 0 according as the proband is a carrier or not and Y_1 is a vector of phenotype indicators, one for each relative. The use of the conditional probability $P(g_0|Y_0)$, instead of $P(g_0,Y_0)$, allows one to preferentially select diseased probands and is justified by assumption A4. The term $P(Y_1|g_0)$, instead of $P(Y_1|g_0,Y_0)$, is based on the conditional independence assumption A3, which also leads to

$$P(Y_1|g_0) = \Sigma P(Y_{11}|g_{11})P(Y_{12}|g_{12}) \ldots$$
$$P(Y_{1m}|g_{1m})P(g_1|g_0), \qquad [2]$$

where the summation is over the vector g_1 of indicators of the m relatives' carrier statuses. In equation 2, $P(Y_{1i}|g_{1i})$ depends only on ϕ_1 for $g_{1i} = 1$ or ϕ_0 for $g_{1i} = 0$. The probability mass function $P(g_1|g_0)$ depends only on the allele frequency $P(A) =$ q and is calculated by standard mendelian methods under HWE. If relatives were also genotyped, the first term in equation 1 would become the product of $P(g_1|g_0)$, which depends only on q, and of $P(Y_1|g_1)$, which depends only on ϕ_1 or ϕ_0. Unless otherwise noted, we shall assume that only the phenotypes of $m = 2$ relatives are available for each proband. Further computational details are available in Gail et al. *(5)* and in a companion paper *(8)*.

RESULTS

Sample Size

The following example is based on estimates by Claus et al. *(6)* of the penetrance of alleles predisposing to breast cancer. Adapting their results for dichotomous outcomes, Gail et al. *(5)* used $q = 0.0033$, $\phi_1 = 0.92$, and $\phi_0 = 0.10$ to represent a rare autosomal dominant allele with high penetrance.

In order to estimate ϕ_1 with precision $\pm1.96\{\mathrm{Var}(\hat{\phi}_1)\}^{1/2} = \pm0.05$, a cohort study would require 114 mutation carriers. To identify these carriers, however, one would expect to need to screen 17 301 subjects because the allele is so rare. This might be possible in the context of a retrospective cohort study in which biologic materials were stored on a large number of women, but the effort would be substantial. A prospective cohort study would require many years of follow-up after the massive screening effort.

It is also possible to estimate ϕ_1 and ϕ_0 from a population-based case–control study if the probability of disease $P(Y_0)$ is known [or if the allele frequency is known, as discussed by Gail et al. *(5)*]. Gail et al. *(5)* found that 17 030 case subjects and control subjects would need to be genotyped to achieve the desired precision on $\hat{\phi}_1$ using the optimal case:control ratio, 1524:15 506.

For comparison, a kin-cohort design in which 10% of the probands are cases and 90% are controls requires genotyping 16 080 probands if there are $m = 1$ relatives per proband. If there are $m = 2$ relatives per proband, 14 935 probands are required *(5)*. Surprisingly, using only case probands does not decrease the required sample sizes necessarily. With one relative per proband, 26 851 case probands are needed, instead of the 16 080 probands above, while with two relatives per proband, 13 418 case probands are needed instead of 14 935.

Substantial reductions in the numbers of required genotypes can be obtained in the kin-cohort design if it is possible to genotype the relatives. For example, with $m = 1$ and with all case probands, if one genotypes both the proband and the relative, then 3549 families are required and $2 \times 3549 = 7098$ genotypes. If one only genotypes the relative when the proband is a carrier, then 3940 families and 4167 genotypes are required *(5)*.

Thus, the kin-cohort design offers only modest reductions in sample size, compared with cohort or case–control designs, unless it is also possible to genotype relatives. Siegmund et al. *(9)* describe increases in efficiency from genotyping additional family members in two-stage designs.

As the allele frequency increases, the required sample sizes for the kin-cohort design decrease dramatically, as is also the case for cohort and case–control designs. Gail et al. *(5)* give tables that allow one to estimate sample sizes required to obtain required precision on estimates of ϕ_1. These tables cover a range of allele frequencies, allow for $m = 1$ or 2 relatives per proband, and allow for the possibility of genotyping a relative.

Violations of Assumptions

Selection bias. A serious bias can arise if probands agree to participate more readily if they have affected relatives, violating assumption A4. Gail et al. *(8)* examined families with a mother proband and $m = 2$ daughter relatives and with $\phi_1 = 0.9$, $\phi_0 = 0.1$, and $q = 0.1$. They assumed that a mother with one affected daughter was twice as likely to participate as a mother with no affected daughters and that a mother with two affected daughters was four times as likely to participate, regardless of the mother's phenotype. (We call this "1:2:4 selection bias.") They used designs with 1154 families to achieve precision ±5% on $\hat{\phi}_1$. In simulations with 5000 repetitions, they found mean estimates (with standard errors) of 0.944 (0.009) for $\hat{\phi}_1$, 0.150 (0.015) for $\hat{\phi}_0$, and 0.205 (0.015) for $\hat{q}$. Thus, both the penetrances and allele frequency were seriously overestimated.

Similar overestimates were found when parameter values $\phi_1 = 0.56$, $\phi_0 = 0.13$, and $q = 0.0116$ were chosen to reflect the data in Struewing et al. *(2)*. Because ϕ_1 is nearer 0.5 than in the previous example, we required more families (i.e., 7634) to achieve the required precision, and half of these families had case probands. The average estimates (with standard errors) were 0.690 (0.021), 0.214 (0.003), and 0.018 (0.001), respectively, based on 559 simulations. Thus, all the parameters were overestimated, as in the previous example.

Cannings and Thompson *(10)* commented on similar ascertainment biases in segregation analyses that arise when families are selected partly on the basis of relatives' phenotypes. Siegmund et al. *(9)* reviewed methods for systematically oversampling probands with affected relatives. When such oversampling is designed, analytical corrections are available to eliminate bias; however, in usual kin-cohort designs, one has no control over who volunteers. Even in kin-cohort designs in which one can select representative samples of probands, one will not control which selected probands agree to participate or know why some refuse.

Misclassification of relatives' phenotypes. If one relies on the proband to provide information on relatives' phenotypes, there is some chance of misclassification. Gail et al. *(8)* studied designs with 5893 families, chosen to achieve ±5% precision on $\phi_1 = 0.9$, with $\phi_0 = 0.1$ and $q = 0.01$. They defined sensitivity as the chance that a diseased relative will be described as diseased by the proband and specificity as the chance that a nondiseased relative will be described as nondiseased. A sensitivity of 0.9 induced a downward bias of about 10% in $\hat{\phi}_0$ and $\hat{q}$ and a downward bias of 3% in $\hat{\phi}_1$. A specificity of 0.9 induced large upward biases; average estimates of $\hat{\phi}_0$, and $\hat{q}$ (with standard errors) in 1000 simulations were 0.945 (0.016), 0.184 (0.004), and 0.017 (0.002), respectively.

Violation of the conditional independence assumption A3. Gail et al. *(8)* examined a logistic model that included a normally distributed familial effect, *b*, with mean zero and variance τ^2. The random familial effect is drawn independently for each family, and each member of a family shares the same random effect. A member of a family with carrier status *g* and familial random effect *b* has a probability of disease

$$P(Y = 1 \mid g, b) = \{1 + \exp(-\mu_g - b)\}^{-1}, \qquad [3]$$

where μ_1 and μ_0 correspond to $g = 1$ and 0 and are chosen so that the marginal probabilities satisfy $P(Y = 1 \mid g = 1) = 0.9$ and $P(Y = 1 \mid g = 0) = 0.1$. For example, for $\tau^2 = 4.0$, $\mu_1 = 3.4095$, and $\mu_0 = -3.4095$.

The random familial effect induces residual correlations among family members' phenotypes, conditional on their genotypes. For example, with $\tau^2 = 4.0$, the correlation between the phenotypes of two family members, each with genotype $g = 1$, is 0.287. To determine the effects of such familial effects, Gail et al. *(8)* simulated 1000 studies, each based on 2178 families, 10% of which had case probands. Mean estimates of ϕ_1, ϕ_0, and q (with standard errors) from the standard model without random effects were 0.948 (0.026), 0.087 (0.005), and 0.015 (0.002), respectively. Thus, the random familial effect leads to overestimation of $P(Y = 1 \mid g = 1)$ and q and underestimation of $P(y = 1 \mid g = 0)$. It is as if estimates of parameters of the genetic model are exaggerated to accommodate the component of familial aggregation not imparted by the gene under study.

We repeated these analyses using parameter estimates based on the data of Struewing et al. *(2)*; namely, $\phi_1 = 0.56$, $\phi_0 = 0.13$, and $q = 0.0116$. Setting $\tau^2 = 4.0$ in the previous random effects model, we required $\mu_1 = 0.3980$ and $\mu_0 = -2.9948$ to preserve the marginal probabilities $P(Y = 1 \mid g = 1) = 0.56$ and $P(Y = 1 \mid g = 0) = 0.13$. Mean estimates of ϕ_1, ϕ_0, and q (with standard errors) from the standard model without random effects were 0.668 (0.046), 0.113 (0.003), and 0.014 (0.001), respectively. Thus, ignoring the random effect increased the estimated penetrance for carriers and the allele frequency, while it decreased the estimated penetrance for noncarriers. These results resemble those in the previous example.

An independently segregating mutant allele, *C*, could also induce residual familial correlation given genotypes for *A* and *a*. Gail et al. *(8)* found that, even if *C* had high penetrance, it induced little bias in estimates of $P(Y = 1 \mid g = 1)$, $P(Y = 1 \mid g = 0)$, and q from a naive model based only on alleles *A* and *a*, provided *C* was uncommon. In contrast, simple segregation analysis of these data would have yielded results that combined the effects of *A* and *C* and estimated a combined frequency of both *A* and *C*. Because the kin-cohort analysis is based on genotyping the proband for *A* and *a* and because a rare allele *C* induces relatively little residual correlation, serious bias is avoided.

Violation of HWE. Gail et al. *(5,8)* studied a population consisting of two noncommunicating strata within each of which mating was random and HWE held but between which there was no mating. Therefore, HWE does not hold in the overall population. The allele frequencies within strata were chosen to preserve the carrier frequency in the entire population. Gail et al. *(5,8)* found little evidence for bias in estimates of ϕ_1, ϕ_0, or q in the presence of such violations of HWE, provided allele frequencies were small, because the joint distribution of Y_1, Y_0, g_1, and g_0 was little affected by such stratification.

Validity of asymptotics. As indicated by the examples in the "Sample Size" section, large samples are needed to achieve good precision for estimates of ϕ_1 when allele frequencies are rare. Fig. 1, taken from Gail et al. *(8)*, shows a histogram of values of $\hat{\phi}_1$ based on 5000 simulations, each with 589 case probands and 5304 control probands and with $m = 2$ relatives (daughters) in each family. These numbers of probands were chosen to achieve ±5% precision on $\hat{\phi}_1$. The distribution of $\hat{\phi}_1$ is symmetric (Fig. 1), and the average estimates (with standard errors) of $\phi_1 = 0.90$, $\phi_0 = 0.10$, and $q = 0.01$ were 0.900 (0.026), 0.100 (0.003), and 0.010 (0.001), respectively. The coverage of a nominal 95% Wald confidence interval for ϕ_1 was 0.932, and that of a profile likelihood-based confidence interval was 0.945.
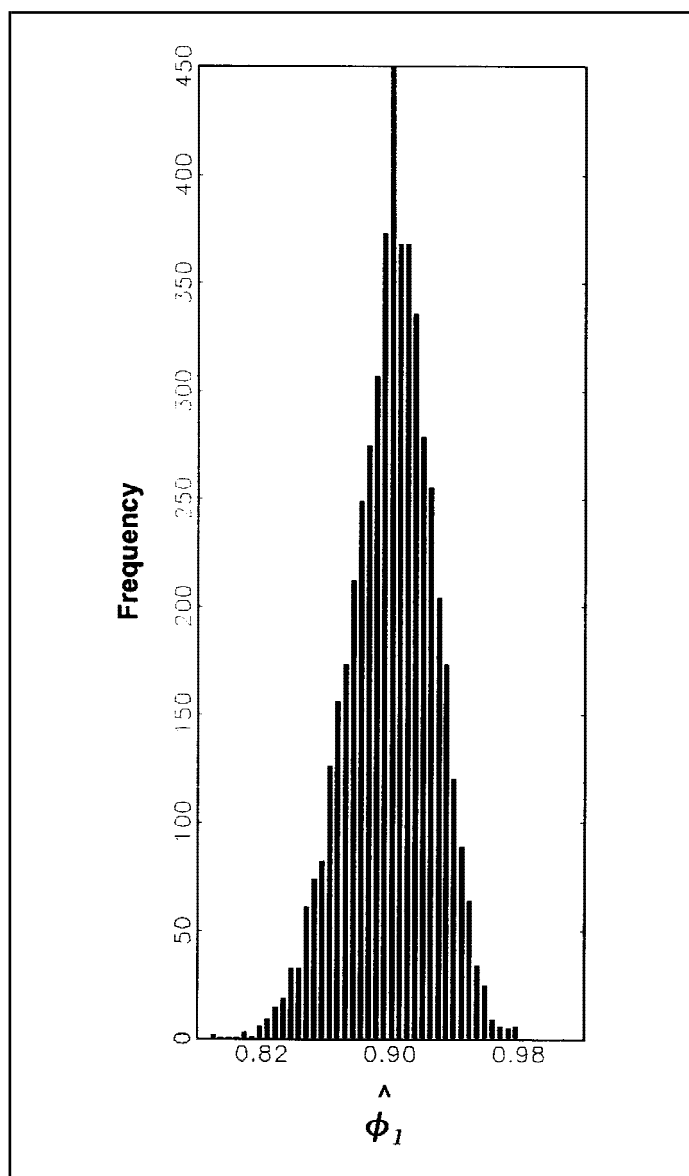
Fig. 1. Histogram of $\hat{\phi}_1$ for samples with 5893 families, 10% of which have case probands. The corresponding parameter values are $\phi_1 = 0.9$, $\phi_0 = 0.1$, and $q = 0.01$.
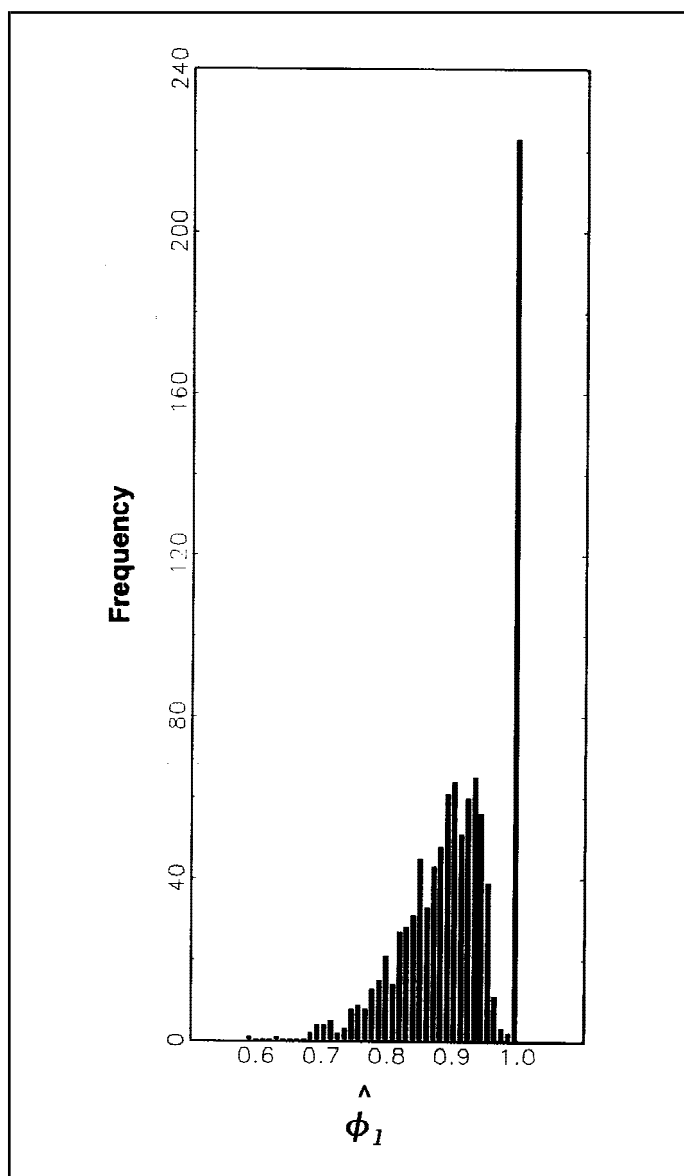

Fig. 2. Histogram of $\hat{\phi}_1$ for samples with 737 families, 10% of which have case probands. The corresponding parameter values are $\phi_1 = 0.9$, $\phi_0 = 0.1$, and $q = 0.01$.

By contrast, a smaller study with 74 case probands and 663 control probands often yielded estimates $\hat{\phi}_1$ on the boundary $\hat{\phi}_1 = 1.0$ (Fig. 2). The small sample distribution of $\hat{\phi}_1$ was distinctly non-normal, and the Wald confidence interval had coverage 0.773, based on 1000 simulations. It is encouraging that the profile likelihood-based confidence interval for ϕ_1 had coverage 0.942.

Survival data. We investigated whether survival data would be subject to similar selection bias and bias from ignoring violations of the conditional independence assumption A3. We assumed that the time to cancer onset among carriers followed a Weibull distribution, $P(T \geq t) = \exp\{-(\lambda t)^\alpha\}$, where $\lambda = 0.013024$ and $\alpha = 2.133435$ were chosen to match the cumulative risk of developing breast cancer to ages 50 and 70 years found by Struewing et al. *(2),* namely, 0.33 and 0.56. Likewise, a Weibull model with $\lambda = 0.007845$ and $\alpha = 3.289313$ matches the cumulative risk to ages 50 and 70 years of noncarriers found by Struewing et al. *(2),* namely, 0.045 and 0.13. The allele

frequency was $q = 0.0116$. The simulated potential ages of the sister proband, sister, and mother were assumed to be multivariate normal with means and covariances chosen to match data in Struewing et al. *(2),* and censoring from competing mortality and from death following breast cancer onset was taken into account as described by Gail et al. *(5).*

We assessed $1:2:4$ selection bias *(see above)* from 10 simulations, each including 10 000 families, of which 25% had case probands. The mean cumulative risk to age 70 years in carriers (with standard error) was 0.705 (0.022), the mean cumulative risk in noncarriers was 0.209 (0.003), and the mean estimated allele frequency was 0.018 (0.002). Thus, as for dichotomous outcomes, $1:2:4$ selection bias led to overestimation of the cumulative risk in carriers and noncarriers and of the allele frequency.

We studied violation of the conditional independence assumption by multiplying all hazards within a given family by a frailty factor, γ, that has a chi-squared distribution with 1 *df.*

Thus, γ has mean 1 and variance 2. Weibull parameters were adjusted to preserve marginal probabilities in the presence of this frailty, of $P(T \leqslant 50 \mid g = 1) = 0.33$, $P(T \leqslant 70 \mid g = 1) = 0.56$, $P(T \leqslant 50 \mid g = 0) = 0.045$, and $P(T \leqslant 70 \mid g = 0) = 0.13$. We conducted 100 simulations each based on 5000 families, half of which had carrier probands. Using the naive survival model without frailty, we observed average cumulative risks (with standard errors) to age 70 of 0.681 (0.060) for carriers and 0.12 (0.003) for noncarriers and an average estimated allele frequency of 0.014 (0.001). Thus, as for dichotomous data, ignoring residual familial correlation led to overestimation of the penetrance for carriers, underestimation of the penetrance for noncarriers, and overestimation of the allele frequency.

DISCUSSION

The kin-cohort design has several practical advantages. In favorable circumstances, this design can be implemented quickly, compared with cohort or population-based case–control designs because information on relatives can be obtained from a single interview of the proband and because one is relying on the previous history of disease in the proband and relatives to estimate penetrance. Moreover, information on several diseases can be obtained from a single proband interview, and several mutations can be studied by appropriate genotyping of the proband. Gail et al. *(5)* found that the kin-cohort design usually requires slightly smaller sample sizes than cohort or population-based case–control designs to achieve the same precision on estimates of the penetrance among mutation carriers, and the reductions in sample size can be considerable if some relatives of the proband can also be genotyped.

Despite the potential practical advantages of the kin-cohort design, it is susceptible to certain biases that affect other forms of segregation analysis. In particular, penetrances and allele frequencies can be seriously overestimated if the probability that a potential proband participates increases with the number of affected relatives of the proband. Two-stage designs that first determine the number of affected relatives before attempting to recruit a proband afford the possibility of bias correction *(9)*. Standard population-based case–control designs and cohort designs are relatively impervious to this type of ascertainment bias *(5)*.

Imperfect specificity in determining the phenotypes of relatives also leads to overestimates of penetrance and allele frequencies. This finding highlights the desirability of applying the kin-cohort design only in populations that can provide good family history data or in special settings with registry data that can provide the needed information on phenotypes of family members *(1)*. Such a study of a founder mutation in BRCA2 in Iceland, where there is a cancer registry and recorded links to family members, yielded estimates of cumulative breast cancer risk to age 70 years of 37% *(11)*.

Residual familial correlations induced by factors such as common dietary habits can lead to an exaggerated estimate of the effects of the gene under study and, in particular, to overestimates of $P(Y = 1 \mid g = 1)$ and q and underestimates of $P(Y = 1 \mid g = 0)$. We found similar results for survival data. On the other hand, another rare gene that segregates independently has little effect on estimates of these parameters for the gene under study in the kin-cohort design.

Violations of HWE have little effect on estimates of penetrance.

Very large samples may be needed to ensure the validity of standard asymptotic approximations. In some examples, however, likelihood-based confidence intervals for ϕ_1 had near-nominal coverage even when the distribution of $\hat{\phi}_1$ was evidently non-normal. This problem can be alleviated by studying special populations with higher allele frequencies because the number of families needed to attain precise estimates and good performance of asymptotic approximations is reduced in such populations.

In many ways, the study of the penetrance of mutations of BRCA1 and BRCA2 in the Washington Ashkenazi study by Struewing et al. *(2)* was an ideal application for the kin-cohort design. Indeed, Wacholder et al. *(1)* developed the design with this study in mind. The prevalence of mutant alleles of about $q = 0.01$ in this population reduced the required sample sizes, compared with a study in the general population with allele frequency 0.003, say. Most of the probands were well educated and could provide good information on the breast cancer status of first-degree relatives. Although there was some evidence that the probands had more affected relatives than in the general population of Ashkenazi women *(2)*, the effect of such an ascertainment bias would be to overestimate penetrance. The striking result of this study, however, was that the estimate of penetrance was lower than that obtained from highly affected pedigrees *(3)*, a result that cannot be explained by such a bias.

An important concern in studies of gene penetrance is the need to respect the autonomy of study participants and to protect them from harm that might result from the improper use or release of confidential genetic information. One issue is whether the genotyped participants will be offered genotype information. This should only be done in studies that provide counseling to determine the wishes of the participants and to convey genetic information appropriately. The institutional review board overseeing the study by Struewing et al. *(2)* determined that genotypes should not be revealed to participants and required that all data linking genotypes with individuals be eliminated after the main analysis. Kin-cohort designs that seek additional interview data or samples from relatives selectively on the basis of the genotype of the proband may inadvertently convey unwanted information on the risk of carrying an adverse mutation, either to the proband or to the relatives. Special care would be needed, therefore, in obtaining informed consent for such designs.

One problem that arose in the context of the study by Struewing et al. *(2)* was that more than one volunteer proband would sometimes come from the same family. In this case, it is not evident how to modify the likelihood *(1)*. Although standard methods can be used to obtain representative samples of individual probands, it is not clear how to sample families without reference to a proband. Indeed, although the set of relatives of a sampled proband is well defined, it is not clear how to define a family without reference to a proband. Thus, there are some conceptual issues that require further clarification for population-based inference when one goes beyond the framework of representative sampling of probands.

In principle, the kin-cohort design has other potential applications *(1)*. One can introduce covariates into the penetrance function and use regression methods to estimate genetic effects, the effects of measured covariates, and their interactions. A potential weakness of this strategy may be in obtaining accurate data on covariates in relatives. The kin-cohort design can also be used to study the main effects of two independently segregating

genes, as well as their interactions. Particularly large samples might be required to test for an interaction if only the proband is genotyped, however, because the genotype of the proband gives only partial information on the genotype of the relatives. Other potential applications include evaluation of the effects of a particular gene on other characteristics of the relative, such as their survival following cancer onset *(12)*. It would seem that a more powerful approach for such studies would be to genotype the relatives rather than to rely on less direct information from the proband's genotype, but ethical or cost constraints may prevent genotyping relatives, and some of them may have died.

There is scope for further methodologic work for survival models. Wacholder et al. *(1)* took a nonparametric approach. They recognized that the survival curve (1 – the cumulative risk function) of a relative of a proband who is a carrier is a mixture (approximately 50 : 50) of survival curves for carriers and noncarriers, whereas the survival curve for a relative of a proband who is a noncarrier is another mixture (approximately 0 : 100) of survival curves for carriers and noncarriers. Using Kaplan–Meier estimates of the observable mixed survival distributions for relatives of carrier probands and for relatives of noncarrier probands, they solved two linear equations for the underlying survival distributions for carriers and noncarriers. Struewing et al. *(2)* were not explicit about how allele frequencies were estimated, and these allele frequencies are needed to define the mixing coefficients. Nonetheless, the solutions are insensitive to the exact allele frequencies for rare alleles. Although the procedures used in the study by Struewing et al. are consistent, provided estimates of allele frequencies are consistent, in small samples they can lead to nonmonotonic estimates of the survival curves for carriers and noncarriers.

In principle, parametric approaches, such as those used by Gail et al. *(5)* for a three-parameter improper Weibull model for carriers and for noncarriers, can be used for inference in a manner similar to that described in the "Survival data" section. In moving toward a nonparametric solution, one could consider piecewise exponential models instead, as used by Claus et al. *(6).* These approaches have the advantage that estimated survival curves will be monotonic, provided parameters are fit subject to their natural constraints. As the number of parameters increases, however, full likelihood methods may become unstable and may not converge, and there is no guarantee of consistency if the number of parameters increases at the same rate as the sample size. Thus, further work is needed to develop survival estimates that are weakly parametric or nonparametric and to allow for potential residual familial correlation.

REFERENCES

(1) Wacholder S, Hartge P, Struewing JP, Pee D, McAdams M, Brody L, et al. The kin-cohort study for estimating penetrance. Am J Epidemiol 1998;148: 623–30.

(2) Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 1997;336:1401–8.

(3) Easton DF, Ford D, Bishop DT. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Am J Hum Genet 1995;56:265–71.

(4) Whittemore AS, Gong G, Itnyre J. Prevalence and contribution of BRCA1 mutations in breast cancer and ovarian cancer: results from three U.S. population-based case–control studies of ovarian cancer. Am J Hum Genet 1997;60:496–504.

(5) Gail MH, Pee D, Benichou J, Carroll R. Designing studies to estimate the penetrance of an identified autosomal dominant mutation: cohort, case–control and genotyped-proband designs. Genet Epidemiol 1999;16:15–39.

(6) Claus EB, Risch NJ, Thompson WD. Genetic analysis of breast cancer in the Cancer and Steroid Hormone Study. Am J Hum Genet 1991;48:232–42.

(7) Thomas DC. Design of gene characterization studies: an overview. Monogr Natl Cancer Inst 1999;26:17–23.

(8) Gail MH, Pee D, Carroll R. Properties of likelihood methods for estimating the penetrance of an autosomal dominant mutation in kin-cohort studies. J Statist Planning Inference. In press 1999.

(9) Siegmund KD, Whittemore AS, Thomas DC. Multistage sampling for disease family registries. Monogr Natl Cancer Inst 1999;26:43–8.

(10) Cannings C, Thompson EA. Ascertainment in the sequential sampling of pedigrees. Clin Genet 1977;12:208–12.

(11) Thorlascius S, Struewing JP, Hartge P, Olafsdottir GH, Sigvaldason H, Tryggvadottir L, et al. Population-based study of risk of breast cancer in carriers of BRCA2 mutation. Lancet 1998;352:1337–9.

(12) Lee JS, Struewing JP, Wacholder S, McAdams M, Pee D, Brody LC, et al. Survival after breast cancer in Ashkenazi Jewish BRCA1/2 mutation carriers. J Natl Cancer Inst 1999;91:259–63.

NOTE